NICE National Institute for Health and Care Excellence

The guidelines manual: appendices J–K

Tailored service improvement support

Published: 30 November 2012
nice.org.uk

Contents

Appendix J: Examples of evidence tables

J1: Example of an evidence table for intervention studies

This table is also suitable for diagnostic studies that compare the effectiveness of two or more tests. This only applies if the test is included as part of a test-and-treat strategy – otherwise the evidence table for studies of diagnostic test accuracy (J2) should be used.

Title: (review question)

Bibliographic reference	Study type	Number of patients	Patient characteristics	Intervention	Comparison	Length of follow-up	Outcome measures and effect size	Source of funding
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
.								

[1] Bibliographic reference: author(s), year, article title, journal, volume, pages.

[2] Study type: for example, randomised controlled trial, cohort or case-control studies.

[3] Number of patients: total number of patients included in the study, including number of patients in each arm, with inclusion and exclusion criteria. Also record the numbers of patients who started and completed the study.

[4] Patient characteristics: characteristics relevant to the area of interest: age, sex, ethnic origin, comorbidity, disease status, community- or hospital-based.

[5] Intervention: treatment, procedure or test studied. If important for the study, specify duration of treatment. For diagnostic studies the intervention is the diagnostic test plus associated treatment studied.

[6] Comparison: placebo or alternative treatment. For diagnostic studies, comparison of the test is with another test and treatment strategy.

[7] Length of follow-up: the length of time that patients take part in the study for, from first staging treatment until either a pre-specified end-point (for example, death, specified length of disease-free remission) or the end of the data-gathering phase is reached. If the study is stopped earlier than originally planned for any reason, this should be noted here.

[8] Outcome measures: list all outcome measures defined in the review protocol, including associated harms. For studies with a diagnostic component there will be two interventions to consider – the diagnostic test used and the associated treatment. Use a separate line for each outcome.

Effect size: for example, raw data from the study that allow analyses such as absolute risk reduction and relative risk (reduction), number needed to treat, number needed to harm, odds ratios, as required. Give confidence intervals whenever possible.

[9] Source of funding: government funding (for example, NHS), voluntary/charity (for example, Wellcome Trust), pharmaceutical company; and the role of funding organisations.

[10] Additional comments: additional characteristics and/or interpretations of the studies that the reviewer wishes to record. These might include important flaws in the study not identifiable from other data in the table, and additional questions or issues that will need to be considered but do not figure in the results tables in the study.

J2: Example of an evidence table for studies of diagnostic test accuracy

Title: (review question)

Bibliographic reference	Study type	Study quality	Number of patients	Prevalence	Patient characteristics	Type of test	Reference standard	Sensitivity and specificity Or raw data for 2 x 2 table	Posi and nega pred valu
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
.									

[1] Bibliographic reference: author(s), year, article title, journal, volume, pages.

[2] Study type: for example, cross-sectional, cohort or case–control studies.

[3] Study quality: note particular strengths and weaknesses.

[4] Number of patients: total number of patients included in the study, with inclusion and exclusion criteria.

[5] Prevalence: proportion of people with the disease in the population at risk.

[6] Patient characteristics: characteristics relevant to the area of interest: age, sex, ethnic origin, comorbidity, disease status, community- or hospital-based.

[7] Type of test: description of the diagnostic test used in the study. Specify the test threshold where applicable.

[8] Reference standard: used as a measure of outcome. Specify if it is a 'gold standard' or 'current best practice'.

[9] Sensitivity: proportion of individuals classified as positive by the gold (or reference) standard who are correctly identified by the study test.
Specificity: proportion of individuals classified as negative by the gold (or reference) standard who are correctly identified by the study test.

Raw data for 2 x 2 table: study data collected from tests to calculate sensitivity, specificity, and positive and negative predictive values (see example table below)

		Disease or outcome	
		Present	Absent
Test	+	a (true positive)	b (false positive)
	–	c (false negative)	d (true negative)

[10] Positive predictive value: proportion of individuals with a positive test result who actually have the disease.
Negative predictive value: proportion of individuals with a negative test result who do not have the disease.

[11] Source of funding: government funding (for example, NHS), voluntary/charity (for example, Wellcome Trust), pharmaceutical company; and the role of funding organisations.

[12] Additional comments: additional characteristics and/or interpretations of the studies that the reviewer wishes to record. These might include important flaws in the study not identifiable from other data in the table, and additional questions or issues that will need to be considered but do not figure in the results tables in the study (for example, if a test is one of a sequence of tests; if its utility was determined).

J3: Example of an evidence table for prognostic studies

Title: (review question)

Bibliographic reference	Study type	Study quality	Number of patients	Patient characteristics	Prognostic factor(s)	Length of follow-up	Outcome measures	Results	Sourc of fundir
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
.									

[1] Bibliographic reference: author(s), year, article title, journal, volume, pages.

[2] Study type: for example, cohort, nested cohort, case series.

[3] Study quality: note particular strengths and weaknesses.

[4] Number of patients: total number of patients included in the study, including number and proportion of patients with prognostic factor(s), with inclusion and exclusion criteria. Also record numbers of patients who started and completed the study.

[5] Patient characteristics: characteristics relevant to the area of interest: age, sex, ethnic origin, comorbidity, disease status, community- or hospital-based. Include method used to select participants.

[6] Prognostic factor(s): include details of method of measurement.

[7] Length of follow-up: the length of time that patients take part in the study for, from entry until either a pre-specified end-point (for example, death, specified length of disease-free remission) or the end of the data-gathering phase is reached. If the study is stopped earlier than originally planned for any reason, this should be noted here.

[8] Outcome measures: all outcome measures should be listed, with each on a separate line.

[9] Results: relative risk or hazard associated with the prognostic factor of interest; absolute risk of event in baseline group; time-to-event analysis.

[10] Source of funding: government funding (for example, NHS), voluntary/charity (for example, Wellcome Trust), pharmaceutical company; and the role of funding organisations.

[11] Additional comments: additional characteristics and/or interpretations of the studies that the reviewer wishes to record. These might include important flaws in the study not identifiable from other data in the table, and additional questions or issues that will need to be considered but do not figure in the results tables in the study.

J4: Example of an evidence table for qualitative studies

Title: (review question)

Reference	Research parameters				Population	Outcomes	Funding	Additiona
Bibliographic reference	Research question	Theoretical approach	Data collection	Method and process of analysis	Population and sample collection	Key themes	Source of funding	Limitation
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
.								

[1] Bibliographic reference: author(s), year, article title, journal, volume, pages.

[2] Research question: what was/were the research question(s)?

[3] Theoretical approach: what theoretical approach (for example, grounded theory, interpretive phenomenological analysis) does the study take (if specified)?

[4] Data collection: how were the data collected? Give details of:

- method(s)
- by whom
- setting(s)
- when.

[5] Method and process of analysis: what methods were used to analyse the data (for example, constant comparative method)?

[6] Population and sample collection: what population was the sample recruited from? Include the following information:

- how they were recruited (for example, specify the type of purposive sampling)
- how many participants were recruited
- specific exclusion criteria
- specific inclusion criteria.

[7] Key themes: list all relevant to this review (with illustrative quotes if available).

[8] Source of funding: government funding (for example, NHS), voluntary/charity (for example, Wellcome Trust), pharmaceutical company; and the role of funding organisations.

[9] Limitations: both those identified by the author(s) and those identified by the reviewer.

[10] Evidence gap and/or recommendations for future research.

Appendix K: GRADE profile and economic evidence profile

K1: Worked example of a GRADE profile

Review question: Should duloxetine vs placebo be used for painful diabetic neuropathy?

Quality assessment							No. of patie
No. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Duloxetine
Patient-reported 30% pain reduction (follow-up 12 weeks)							
2[1]	Randomised trials	No serious risk of bias	Serious[2]	No serious indirectness	No serious imprecision	None	220/327
Patient-reported 50% pain reduction (follow-up 12 weeks)							
4[3]	Randomised trials	No serious risk of bias	Serious[4]	No serious indirectness	Serious imprecision[5]	None	485/896
No. of withdrawals due to adverse effects (follow-up 12 weeks)							
4[3]	Randomised trials	No serious risk of bias	No serious inconsistency	No serious indirectness	No serious imprecision	None	113/906
Dizziness (adverse effects) (follow-up 12 weeks)							

3[6]	Randomised trials	No serious risk of bias	No serious inconsistency	No serious indirectness	Serious imprecision[5]	None	90/674
Dry mouth (adverse effects) (follow-up 12 weeks)							
2[7]	Randomised trials	No serious risk of bias	No serious inconsistency	No serious indirectness	Serious imprecision[5]	None	37/448
GI disturbances (adverse effects) (follow-up 12 weeks)							
2[8]	Randomised trials	No serious risk of bias	No serious inconsistency	No serious indirectness	Serious imprecision[5]	None	28/332
Any adverse effects (non-specified) (follow-up 12 weeks)							
1[9]	Randomised trials	No serious risk of bias	No serious inconsistency	No serious indirectness	Very serious imprecision[10]	None	86/106

[1] Gao et al. (2010); Wernicke et al. (2006).

[2] Substantial heterogeneity, random-effect model was used. Potential sources of heterogeneity: i) Gao et
between 30 mg and 120 mg, non-pharmaceutical company funded; ii) Wernicke et al. (2006) – only per-pr
(60 mg and 120 mg), pharmaceutical company funded.

[3] Gao et al. (2010); Goldstein et al. (2005); Raskin et al. (2005); Wernicke et al. (2006).

[4] Substantial heterogeneity, random-effect model was used. Potential sources of heterogeneity: i) Gao et
120 mg, non-pharmaceutical company funded; ii) Goldstein et al. (2005), Raskin et al. (2005) and Wernick
(20 mg, 60 mg and 120 mg), pharmaceutical company funded.

[5] Confidence interval crossed one end of default MID.

[6]Gao et al. (2010); Goldstein et al. (2005); Wernicke et al. (2006).

[7] Gao et al. (2010); Goldstein et al. (2005).

[8] Gao et al. (2010); Wernicke et al. (2006).

[9] Gao et al. (2010).

[10] Confidence interval crossed both ends of default MID.

Abbreviations: CI, confidence interval; GI, gastrointestinal; ITT, intention to treat; MID, minimal importan

K2: Example of an uncompleted GRADE profile

Quality assessment							No. of patients		Effect	
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	.	.	Relative (95% CI)	Ab
X										
X										
X										
X										
X										
X										
[References, abbreviations and other footnotes].										

K3: Worked example of an economic evidence profile

Adapted from Crohn's disease: management in adults, children and young people (NICE clinical guideline 152).

Systematic review of economic evaluations of budesonide for maintenance of remission in Crohn's disease

Study	Limitations	Applicability	Other comments	Incremental			Uncertainty
				Costs	Effects	ICER	
Noble 1998 Budesonide CIR versus no maintenance therapy	Potentially serious limitations[1,2]	Partially applicable[3]	Study employed a Markov decision-analytic model with a 1-year time horizon	£115	0.017 QALYs[5]	£6,981 per QALY gained	ICER decreases significantly if the cost of surgery is increased.

NCGC model Oral budesonide versus no maintenance therapy[4]	Potentially serious limitations[2]	Directly applicable	Study employed a Markov decision-analytic model with a 2-year time horizon	£477[f] £150[7] £528[8] £336[9]	0.012 QALYs[6] 0.012 QALYs[7] 0.006 QALYs[8] 0.005 QALYs[9]	£40,392 per QALY gained[6] £15,070 per QALY gained[7] £87,610 per QALY gained[8] £65,013 per QALY gained[9]	No treatment most cost-effective option when baseline risk of relapse decreased. In the PSA, probability of budesonide being the most cost-effective treatment at willingness-to-pay threshold of £20,000 per QALY gained ranged from 0 to 8%

[1] Modelling was undertaken over a short time horizon and no probabilistic sensitivity analysis was conducted.

[2] Specific costs and disutilities of drug-related adverse events could not be explicitly modelled. Adverse events were captured by modelling treatment-specific withdrawal rates. This may have overestimated the cost effectiveness of maintenance treatment.

[3] The cost-effectiveness model was designed to reflect the management of Crohn's disease in the Swedish healthcare setting. Although a cost per QALY estimate was reported, it was not based on health-related quality of life values elicited from patients.

[4] The NCGC model compared a number of different maintenance treatments.

[5] Figures may differ because of rounding off.

[6] Conservative 4-line model. Conservative treatment effects were used and people relapsing while on azathioprine maintenance treatment had a different induction sequence.

[7] Conservative 3-line model. Conservative treatment effects were used and people were assumed to have the same induction sequence regardless of maintenance treatment.

[8] Non-conservative 4-line model. Non-conservative treatment effects were used and people relapsing while on azathioprine maintenance treatment had a different induction sequence.

[9] Non-conservative 3-line model. Conservative treatment effects were used and people were assumed to have the same induction sequence regardless of maintenance treatment.

K4: Example of an uncompleted economic evidence profile

Study	Limitations	Applicability	Other comments	Incremental			Uncertainty
				Costs	Effects	ICER	
.							
.							
.							
.							
.							
[References, abbreviations and other footnotes].							